VIOPS03 ライブドア

(株)ライブドア
池邉 智洋
2009年5月29日

www.livedoor.com

livedoor

自己紹介

- ライブドアについて
  - メディア(ポータルサイト)事業
  - データセンター事業
- 私について
  - (株)ライブドア 執行役員
  - メディア事業の開発を担当しています
  - Webプログラマ = 仮想化のエンドユーザー

www.livedoor.com

livedoor

サーバ仮想化

www.livedoor.com

livedoor

仮想化への取組み

- 仮想化は始めたばかり
  (2008年10月～)
- 使用環境 – Xen
  ParaVirtualization
- 仮想化ホスト = まだ100ホスト程度
  → 全体の 5% 位

www.livedoor.com

livedoor

ライブドアのインフラ

- 基本的なインフラの方針
  - 高価なマシンは使わない(使えない)
  - オープンソースを利用
  - スケールアップよりスケールアウト
  - 24/365 のフルマネージドホスティング

www.livedoor.com

livedoor

仮想化に求める事

- サーバの集約/単純化
  - Webサービスの特性に応じた分割
  - 高価なサーバでは無いので密度は低め
  - 1インスタンスの役目は1つに限定
- Deploy/Migration の容易さ
  - カジュアルにサーバを増減

www.livedoor.com

livedoor

集約の組み合わせ

- Webサービスで使用するサーバ
  - www (ReversePROXY)
  - app (mod_perl)
  - MySQL
  - Memcached
  - Batch/MessageQueue etc…
- 消費するリソースによる組み合わせ

www.livedoor.com

livedoor

実際の事例

- livedoor ニュース(1億3000万PV/Month)
- 物理サーバ 50台 => 20台へ
- 仮想化により www/memcached 等のサーバを集約
- MySQL は仮想化せずに運用
  → パフォーマンス重視

www.livedoor.com

livedoor

メリット/デメリット

- デプロイの簡略化
  - アプリ設定済のイメージを保存
  - プログラマの手間は減少
- ホストの管理が煩雑に
  - 1インスタンス = 1マシンの常識からの脱却
  - ホスト管理システムの改修が進行中

www.livedoor.com

livedoor

ストレージ仮想化

www.livedoor.com

livedoor

ストレージの仮想化

- ユーザー参加型サービス中心
  - データは無限に増加
- 巨大なストレージプールが必要
- 一般的にストレージ装置は高価だが
- Webサービスが求める機能に限ると、、

www.livedoor.com

livedoor

ストレージに求める事

- 安価に容量を拡大出来る
- クライアントからは一つに見える
- 高可用性 = データのレプリケーション
- Webのリクエストパターン Read が殆ど
- 過剰な機能/信頼性より価格!

こういう製品はなかなか無い

livedoor

Webサービスのストレージ

- SixApart
  - MogileFS
    http://www.danga.com/mogilefs/
- Facebook
  - Haystack
    http://www.facebook.com/note.php?note_id=76191543919&ref=mf

Webサービスに特化し自作する事で安価に

www.livedoor.com

livedoor

CAP定理

- **C**onsistency
- **A**vailability
- **P**artition Tolerance

3つ全ては満たせないという話

www.livedoor.com

livedoor

Eventually Consistency

- A/P は Webサービスには必須
- Consistency(一貫性)を妥協する
- ストレージの場合
  - 複数ノード間のデータの一貫性
  - ある程度の時間の許容

www.livedoor.com

livedoor

独自実装への道

- 利用シーンを限定する事により実装を簡単に
- Read/Write の割合
  - Read のパフォーマンスは拘るが Write はある程度許容出来る
- 追加/更新/削除
  - 追加が殆んど
- 認証/アクセス制御
  - アプリケーションで制御するのでストレージ側では考慮せず

www.livedoor.com

livedoor

独自実装への道

- オープンソースを利用
- プロトコル
  - Apache を利用する
- データベース
  - MySQL + memcache
- MessageQueue
  - ActiveMQ or Q4M

www.livedoor.com

livedoor

ライブドアでの実装

- ブログ、写真共有等をターゲットにした分散ストレージを Apache モジュール等で実装
- クライアントからは HTTP PUT/GET で操作
- ファイル = URL
  アプリケーションでのパス ⇔URL変換が不要
- ストレージノードは安価なサーバにHDDを積んで使用
- ファイルは複数のノード間にレプリカが作成される
- レプリカの状態がおかしくなったら自動復旧

livedoor

実装

- mod_stf.c
  - データベースを参照してストレージノードを決定して Reverse PROXY
- mod_stf_storage.c
  - PUTメソッドでファイル書き込み
- Worker
  - MessageQueue と連携
  - レプリケーションの作成/復旧etc..

www.livedoor.com

livedoor

Web IF

STF Web Interface

Storage Nodes
Buckets

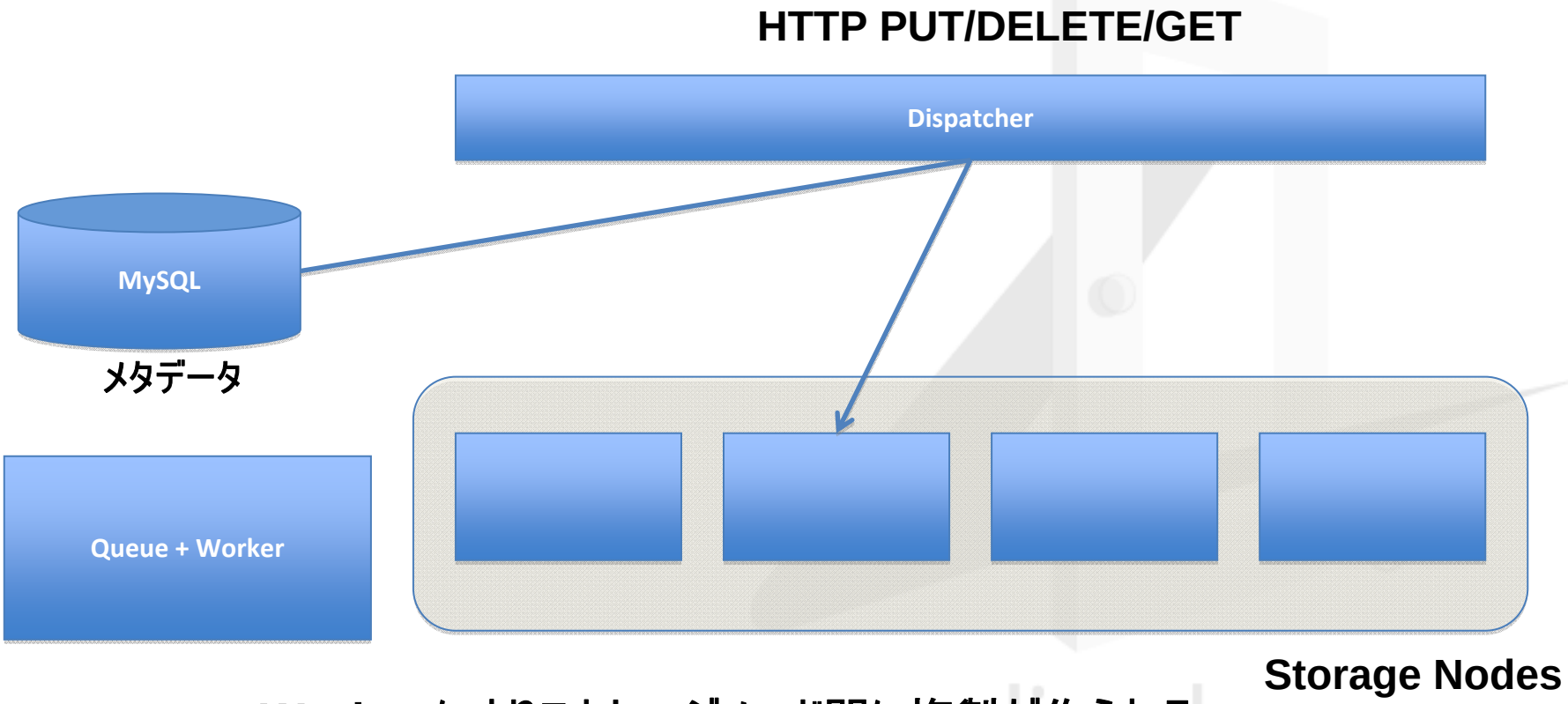
ID	URI	使用領域	DISK容量	Mode
1	http://10.0	90.8G	400.0G	rw
2	http://10.0	91.1G	400.0G	rw
3	http://10.0	91.3G	400.0G	rw
4	http://10.0	91.2G	400.0G	rw
5	http://10.0	91.1G	400.0G	rw
6	http://10.0	91.2G	400.0G	rw
7	http://10.0.	90.9G	400.0G	rw
8	http://10.0.	90.7G	400.0G	rw

ストレージの追加

© 2009 livedoor Co.,Ltd.

livedoor

ストレージ概要



Worker によりストレージノード間に複製が作られる

www.livedoor.com

livedoor

まとめ

- サーバ仮想化は同一構成を並べるWebサービスには有効
- ストレージは要件が特殊
  - 現時点では独自実装
  - オープンソースで公開したいです

www.livedoor.com

livedoor